

Susanne Dobratz
Humboldt-University Berlin
Electronic Publishing Group
University Library/ Computing Centre
dobratz@rz.hu-berlin.de

Panel „Long-term e-Archiving“ (Friday)

At the 2nd Workshop on the Open Archives Initiative: Gaining Independence with e-Print archives and OAI, CERN, Geneva, 17-19 October 2002

Discussion Guidelines

Looking at the motto of this workshop “Gaining independence with e-Print archives and OAI” it suggests first of all that using e-Print publishing methods especially in the sense of a scholarly non-profit publishing independently from any commercial publishing house offers a unique chance to scientists.

A chance to publish their research results:

- Faster
- Cheaper
- More focused on the article’s particular audience.

But the newly won independence from commercial publishing houses demands also a greater responsibility from the scientists. Firstly, the responsibility to find new methods to sell or distribute their scientific output and secondly the responsibility to ensure a long term availability of the published results.

However, it is not only the author that has to take on new responsibilities while pursuing this new opportunity. Following an alternative publishing process, all participants are asked to contribute.

Creator:

- Usage of document formats (text processing system) in order to produce originals, that do not require conversion before put into the e-Print archive.
- Clarify copyright, retain right to e-Print archive
- Use digital signatures

Editor/Reviewer:

- Appropriate quality control: content and formal issues
- Define distribution channels

e-Print Repository Administrator:

- Ensure trustworthiness of repository through
 - Technical security issues
 - Following technical standards like the “Open Archival Information System” model as functional model for an trusted repository architecture, or standard metadata for archiving like Dublin Core, METS, EAD, use standards transport protocols within the archive and in communication with the outside environment
 - Administrative Responsibility
 - Organizational Viability
 - Financial Sustainability
 - Technological and Procedural Suitability

- System Security
- Procedural Accountability

Reader:

- Appropriate use of open access material
- Follow links and references

The emergence of e-Print repositories and the increased availability of digital objects also demands the development of national preservation policies, as proposed in the UK, Australia and the US. A national preservation policy would regulate the legal deposit, the certification of trusted repositories, responsibilities, etc.

How does the OAI movement fit into those thoughts?

Summary

14 people attended the session:

- CLAIVAZ, Jean-Blaise, Librarian, Université de Genève, GENEVA, SWITZERLAND, jean-blaise.claivaz@adm.unige.ch
- CLARA, Kristine, Head of Library & Archives, ITU, GENEVA, SWITZERLAND, kristine.clara@itu.int
- DOBRATZ, Susanne, Head Electronic Publishing Group, Humboldt-University Berlin, BERLIN, GERMANY, dobratz@rz.hu-berlin.de
- GOOVAERTS, Marc, Information Technology staff, Limburgs Uni. Centrum-Library, DIEPENBEEK, BELGIUM, marc.goovaerts@luc.ac.be
- GEWIRTZ, David, Project Manager, Yale University, NEW HAVEN, USA, david.gewirtz@yale.edu
- GROZEL, Gerard, Head of UC DIST Versailles, INRA, VERSAILLES, FRANCE, grozel@versailles.inra.fr
- BETTERMANN, Henrik, Founder, Scientific African e.V., HERDECKE, GERMANY, hbetter@scientific-african.de
- TOWE, Matthias, Librarian, University Library Basel, BASEL, SWITZERLAND, matthias.toewe@unibas.ch
- SIMPSON, Pauline, Head of information services, University of Southampton, SOUTHAMPTON, UK, ps@soc.soton.ac.uk
- NIXON, William, Deputy Head of IT services, Library, University of Glasgow, GLASGOW, UK, w.j.nixon@lib.gla.ac.uk
- VAN ZUNDERT, Joris, Researcher, NIWI, AMSTERDAM, THE NETHERLANDS, joris.van.zundert@niwi.knaw.nl
- CAMERINO, Manuela, Librarian, Istituto Uni. di Scienze Motorie, ROME, ITALY, biblioteca@iusm.it
- PINFIELD, Stephen, Assistant director of inf. Services, University of Nottingham, NOTTINGHAM, UK, stephen.pinfield@nottingham.ac.uk
- FLACK, Howard, Responsible of courses, University of Geneva, GENEVA, SWITZERLAND, Howard.Flack@cryst.unige.ch
- MAHONEY, John, Project Officer, ILRT, University of Bristol, BRISTOL, UK, john.mahoney@bris.ac.uk

Within the discussion nearly all aspects of long-term preservation were mentioned: starting from metadata and documents formats via policies and responsibilities up to security issues and good examples. David Gewirtz, Yale University, reported on a long-term preservation projects for e-journals with Elsevier Science. Funding for the project came from the Mellon Foundation. The project lasted about 18 months and used the OAIS and OAI models to build a prototype digital archive for e-journals. We mentioned the LOCKSS Initiative and discussed the scalability problem with that.

The group shared some common thoughts together:

There is an urgent demand to archive now! But in order to do so, best practise solutions are necessary and digital toolkits available to solve at least a few of the problems that appear by using well-packaged solutions. We raised questions like 'to which format should the digital objects be converted'. There is not yet a definitive long term preservation format answer and it depended on the media type but the recommended format at present for document types was SGML, XML. Conversion tools are needed to carry out this process. Mentioned as an example was the work of Yale, Harvard and PubMed Central to create an archive DTD to

normalize content for e-journals.. At present e-Print Archives advertise open access and do not emphasize preservation activities.

We felt that this emphasis did not support advocacy and that **Long-term preservation** should become a main goal of the e-Prints and e-Theses community. It was noted that Projects like the one from Yale or the DSpace Project at MIT who are taking responsibility for preservation should impact on the e-Prints community.

Long- term preservation of digital objects is a global problem not just for the e-Print community . We need: more Standards for interoperability; best practise solutions; software packages and tools! We need national initiatives, with shared responsibilities among several players to take responsibility for long term preservation based on a network of trusted repositories to ensure that even if an archiving institution does not exist in the future, its' digital archive will be preserved within a national depository.

Literature

1. Trusted Digital Repositories: Attributes and Responsibilities, an RLG-OCLC-Report. Mountain View, CA: RLG, May 2002. <http://www.rlg.org/longterm/repositories.pdf>
2. Metadata Encoding and Transmission Standard (METS), <http://www.loc.gov/standards/mets/>
3. Rebecca Guenther: The joint work of the OCLC-RLG Preservation Metadata Working Group, RLG Open Forum at ALA June 16, 2002, <http://www.rlg.org/longterm/forum02/guenther.html>
4. Cedars Guide To : Preservation Metadata (March 2002), <http://www.leeds.ac.uk/cedars/guideto/metadata/>
5. DSpace System at MIT: <http://dspace.org/index.html>
6. Yale University: Report to the Mellon Foundation: <http://www.library.yale.edu/~okerson/yea/>.
7. National Preservation Office: Building Blocks for a Preservation Policy; Jan. 2001; 28.06.2002 <http://www.bl.uk/npo/>
8. Jones, M. and Beagrie, N.: Preservation Management of Digital Materials: A Handbook; London, The British Library 2001
9. CCSDS (Consultative Committee for Space Data Systems: Reference Model for an Open Archival Information System (OAIS). Red Book, Version July 2001 www.ccsds.org/documents/pdf/CCSDS-650.0-R-2.pdf